

Polices Unicode pour cunéiforme

Paléo-babylonien (cursif et monumental), néo-assyrien, hittite

I. Standard Unicode 5.0 et Cunéiforme: principes généraux et remarques critiques

II. Polices cunéiformes: paléo-babylonien, néo-assyrien, hittite

Installation

Liste et notes

Utilisation: texte automatique

I. STANDARD UNICODE 5.0 ET CUNÉIFORME

Le *Standard Unicode* version 5.0 (juillet 2006, voir <http://unicode.org>) inclut désormais les blocs *Cuneiform* (12000-1236E) et *Cuneiform Numbers and Punctuation* (12400-12462 et 12470-12473).

Unicode est une norme informatique qui spécifie, pour chaque caractère, un numéro unique (point de code), quelle que soit la plate-forme, quel que soit le logiciel et quelle que soit la langue. Cette norme, développée pour nombre de système d'écriture, manquait jusqu'à présent pour le cunéiforme. L'*Initiative for Cuneiform Encoding* (ICE <http://www.jhu.edu/ice/>), fondée à Baltimore en 2000, visait à combler ce manque. En 2006, une proposition finale pour le système d'écriture cunéiforme a été élaborée par Steve Tinney, Michael Everson et Karljürgen Feuerherm (<http://std.dkuug.dk/jtc1/sc2/wg2/docs/n2786.pdf>).

La définition de ces nouveaux blocs vise à combler à un besoin auquel ne pouvaient répondre les polices traditionnelles, non Unicode. Limitées à 256 caractères, elles ne permettaient pas l'intégration des signes au sein d'une même police. Deux, trois, quatre (voire plus) polices étaient nécessaires pour arriver à couvrir l'ensemble des signes paléo-babyloniens, hittites ou néo-assyriens. L'encodage lui-même variait d'une police à l'autre, d'une plate-forme à l'autre. De ce point de vue, la proposition Unicode pour le système cunéiforme devait permettre un encodage plus simple des signes cunéiformes, ainsi que leur utilisation multiplate-forme. Dans l'ensemble, le résultat est plus que positif et l'on

entrevoit les avantages multiples qui peuvent en découler: constitution de corpus électroniques, de bases de données lexicographiques, recherches sur corpus (tri, occurrences, indexation etc., cf. notamment <http://www.jhu.edu/ice/>).




Comme toute innovation, s'il répond en majeure partie aux exigences, le *Standard Unicode 5.0* présente également quelques difficultés et problèmes, qui demanderont certainement des amendements. Les principales critiques ont été formulées par R. Borger dans l'introduction à *List of Neo-Assyrian Cuneiform Signs. A practical and critical guide to the Unicode blocks «Cuneiform» and «Cuneiform Numbers» of Unicode Standard Version 5.0* (<http://www.sumerisches-glossar.de/download/SignListNeoAssyrian.pdf>), compilée par M. Studt (liste révisée de D. Bachmann). Voir également l'article Unicode cuneiform (2007, August 24): *Wikipedia, The Free Encyclopedia* (http://en.wikipedia.org/wiki/Unicode_cuneiform).

Les principes généraux qui ont présidé à l'encodage sont exposés dans le texte de la proposition finale.


1) l'inventaire des caractères repose sur la liste des signes de la période de Ur III (2112-2095), compilée par la *Cuneiform Digital Library Initiative*. Il s'agit là d'une première phase dans la définition des points de code; d'autres phases devraient s'attacher à d'autres époques (Old Akkadien, 2334-2154 et Early Dynastic, 2900-2335, Archaic Cuneiform, jusque 2901).

Ainsi que le souligne R. Borger, la liste proposée est limitée, en ce sens qu'elle ne tient pas compte des différentes époques et régions utilisant le système cunéiforme. Ceci a pour conséquence qu'une série de signes ne reçoivent pas de points de code en hittite par exemple, ou même en néo-assyrien (cf. *infra*). Bon nombre de points de code devraient en outre être ajoutés par la suite pour arriver à une liste complète, couvrant toutes les facettes chronologiques et géographiques du système d'écriture cunéiforme.


2) signes cunéiformes et caractères cunéiformes ne se correspondent pas forcément; signes complexes et signes composés sont distingués (voir Final Proposal, p.7):


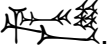
° un signe *complexe* est constitué d'un premier signe dans lequel est inséré un autre signe; l'ensemble forme une unité. Exemple: le signe simple KA  peut servir à former des signes complexes KAxGU  (où GU est inséré dans le premier signe KA), KAxLI , etc. Le signe complexe formant une unité, il reçoit un point de code particulier. Dans les exemples, KA = U12157, KAxGU = U1216B, KAxLI = U12172.

Une série de signes de la liste de Borger (*MeZL*) ne reçoivent pas de points de code. Il s'agit de signes rares: KAxTU, KAxḪAR, LAGABxGI etc., qui représentent en tout bonne trentaine de signes. D'autres sont des signes attestés en hittite (KAxÀŠ, KAxÚR, KAxGAG, KAxGIŠ, KAxPA, KAxLUM, SIxSÁ, EZENxŠE, AMARxKU₆, ÁBxA, KISIM₅xÚ-MAŠ).

° un signe *composé* est constitué de deux ou plusieurs signes organisés en séquence linéaire. Chacun des signes existe en tant que caractère propre. L'ensemble est généralement vu comme une unité, mais chaque composante sera encodée séparément. Exemple: IDIGNA, composé de MAŠ+GÚ+GÀR, c'est-à-dire U12226+U12118+U120FC


La liste du *Standard Unicode 5.0* est ordonnée selon l'ordre de l'alphabet latin, fournissant une description «étymologique» des signes, simples, composés et complexes (cf. Cuneiform Unicode, Wikipedia).

Sur ce plan, souligne encore R. Borger, plutôt que se baser sur les lectures conventionnelles utilisées par les assyriologues, cette liste introduit de nouvelles lectures, dont le but est de décrire le signe (Cuneiform sign A, Cuneiform sign KA times LI, etc.), mais dont l'usage sera souvent peu pratique (il faudra savoir que GIR = ḪAgûnu, que AMAŠ = DAG KISIM5 times LU + plus MASH2). La critique la plus sévère faite à l'encontre de ce nouvel usage concerne d'une part l'analyse des signes et d'autre part l'utilisation parfois laborieuse des points de code. En effet, la «décomposition» de certains signes apparaît comme artificielle dans certains cas et pourrait être discutée: voir, par exemple, Borger, *MeZL*, n° 754 MEŠ, qu'il faudra décomposer en ME+U+U+U (ME = U12228; U+U+U = U1230D) ou n° 459 DUL décomposé en U+TÚG (U1230B+U12306) selon le *Standard Unicode*; selon Borger toutefois, les signes ne sont pas réellement décomposables en deux éléments. Parfois ailleurs, il faudra combiner un, deux, voire trois points de code: U12226+U12118+U120FC pour IDIGNA, U12263+U121EC pour TÛR (NUN-LAGAR), etc. Les ligatures sont également rendues impossibles théoriquement, sauf à constituer une seconde police contenant des variantes de signes spécifiques (qui pourraient être tout à fait artificielles; voir, pour le hittite, les notes à la liste des signes). Pour le néo-assyrien et le paléo-babylonien NIGIN (*MeZL*, n°804) = LAGAB-LAGAB (U121B8+ U121B8), on obtiendra , sans possibilité de rendre réellement la ligature.

Dans d'autres cas, la décomposition de signes peut être problématique; ainsi pour le hittite (1570-1220) qui fait usage d'un syllabaire reposant sur une cursive paléo-babylonienne du nord de la Syrie (E. Neu - C. Rüster, *Hethitisches Zeichenlexikon*, Wiesbaden, 1989). Le signe IDIGNA tel qu'il est usité dans le syllabaire hittite est difficilement décomposable en ses divers éléments: U12226+U12118+U120FC aboutit à , peu ressemblant au signe IDIGNA en tant que tel .

Pour ce type signe, composé, nous avons pris le parti d'en isoler artificiellement les composantes, c'est-à-dire de dessiner une variante de MAŠ, de GÚ, de GÀR, qui permettront d'obtenir le signe final. Ce type de variante est signalée en note dans la liste. L'utilisation de texte automatique permettra également de contourner ce genre de problème (voir ci-dessous).

II. POLICES CUNÉIFORMES: PALÉO-BABYLONIEN, NÉO-ASSYRIEN, HITTITE

Les polices Unicode cunéiforme (TTF) ont été créées par Sylvie Vanséveren.

- Santakku: paléo-babylonien, écriture cursive
- SantakkuM: paléo-babylonien, écriture monumentale (Hammurabi)
- Assurbanipal: néo-assyrien
- Ullikummi (A, B, C): hittite.

Elles sont mises à disposition de la communauté scientifique gratuitement, et ne peuvent être modifiées ou vendues.

Installation des polices

°MacOSX: glisser la (les) police(s) dans User/Library (Bibliothèque)/Fonts.

°Windows (XP):

cliquer sur Démarrer puis sur Panneau de configuration

cliquer sur Polices: le dossier des polices s'ouvre

deux méthodes permettent d'ajouter des polices dans Windows:

- soit copier directement la police dans le dossier (par glisser-déposer)
- soit cliquer sur le menu Fichier de la fenêtre Polices et choisir la commande Installer une nouvelle police. Sélectionner la police dans le champ Liste des polices. Cliquer sur OK.

Liste des signes et notes

On trouvera pour chaque police une liste de signes comprenant les éléments suivants:

- numéro MeZL
- point de code Unicode
- signe (lecture conventionnelle et description Unicode)
- signe dans la police concernée (et variantes pour le hittite)
- autotext

Des notes explicatives et des remarques complètent chaque liste. On y décrit les problèmes spécifiques à un signe, à la décomposition de signes, et aux variantes nécessaires (surtout pour le hittite) à la composition de certains signes complexes et composés.

Texte automatique

Une fois l'encodage des polices réalisé, il reste à pouvoir les utiliser concrètement, c'est-à-dire, à pouvoir écrire en cunéiforme. Il est toujours possible d'accéder aux signes via la Palette de caractères dans MacOSX ou via l'insertion de caractères spéciaux dans les éditeurs de textes, mais cette manière de procéder rend l'écriture extrêmement fastidieuse, puisqu'il faut connaître ou reconnaître chaque signe, ou connaître son point de code spécifique. Pour tenter de faciliter les choses, on pourra recourir à du texte automatique.

L'insertion automatique de texte est possible à réaliser dans des traitements de texte comme Word et Nisus Writer Pro. Des modèles de document sont fournis pour Word. **Chaque modèle est modifiable à loisir par les utilisateurs, qui peuvent ajouter, modifier, supprimer des entrées.**

NB. Neo- et OpenOffice (Mac et PC) ne semblent pas intégrer correctement les polices (ni via l'insertion de caractères spéciaux, ni via la Palette de caractères dans OSX). D'autres traitements (comme Mellel, Mac) ne proposent pas le texte automatique.

Le principe consiste à associer une entrée d'insertion automatique à un signe ou un groupe de signe. Il est ainsi possible d'écrire en cunéiforme sans avoir à connaître les points de code des signes. A cet effet, une série de caractères latins sont également encodés dans ces polices (issus de la police Cunéitruetype, avec l'aimable autorisation du Prof. Dominique Charpin).

Ainsi, en tapant le texte d'entrée "ma" puis en demandant une insertion automatique (via menu ou raccourci clavier, cf. ci-dessous), on obtient le signe voulu (U12220)

𒌷 (Assurbanipal)

𒌷 (Santakku)

𒌷 (SantakkuM)

𒌷 (UllikummiA)

De même, l'entrée "lugal" donnera (U12217)

𒌷 (Assurbanipal)

𒌷 (Santakku)

𒌷 (SantakkuM)

𒌷 (UllikummiA)

Pour les signes composés, faits de plusieurs signes pour chacun desquels un point de code est attribué, le principe permet d'associer les signes:

"mac" (MAŠ)	U12226	𒌷	𒌷
"gu2" (GÚ)	U12118	𒌷	𒌷
"gar3" (GAR)	U120FC	𒌷	𒌷

"idigna" 𒌷𒌷𒌷 (Assurbanipal)
𒌷𒌷𒌷 (SantakkuM)

NB. Pour le paléo-babylonien et le néo-assyrien, un seul modèle de document est proposé (Cuneiform.dot); les entrées de texte automatique sont identiques. Le hittite connaissant des valeurs différentes pour les signes, un autre modèle de document (Hittite.dot) reprend les insertions basées sur les valeurs du hittite.

Les variantes sont marquées par "v" ou "vv"

Assurbanipal: me3 𒌷 me3v 𒌷

Ullikummi: mah 𒌷 mahv 𒌷 mahvv 𒌷

MICROSOFT WORD

(Pour plus d'informations, consulter l'aide de Word ou

<http://support.microsoft.com/kb/207925/fr>)

- **Modèle de document:** Cuneiform.dot, Hittite.dot

- **Installation:**

°MacOSX: Applications\dossier Microsoft Office \Modèles (ou dans le dossier qui contient les modèles personnalisés)

°Windows (XP): C:\Windows\Documents and Settings\User\Application Data\Microsoft\Modèles (ce dossier est caché par défaut).

- **Choisir le modèle** via la Bibliothèque de projets.

Utilisation des insertions automatiques

Il existe différentes façons de procéder:

- soit passer par le menu Insertion → Insertion automatique. Une liste déroulante donne accès aux différentes insertions.
- soit utiliser un raccourci clavier spécifique à l'insertion automatique.

Pour accéder à la boîte de dialogue de définition des insertions automatiques:

- menu Insertion → Insertion automatique (haut de la liste déroulante)
- ou cliquer sur le bouton "insertion automatique"
- ou via le menu Outils → Correction automatique

Pour définir un raccourci clavier d'insertion automatique

Menu Outils → Personnaliser le clavier

Dans la boîte de dialogue, sélectionner à gauche Toutes les commandes

Choisir InsertionAutomatique et définir le raccourci clavier.

Pour avoir accès au bouton "insertion automatique"

Menu Outils → Personnaliser les barres d'outils/menus

Faire glisser le bouton dans une barre d'outil (la barre d'outils standard par exemple).

NISUS WRITER PRO (MAC uniquement)

Dans Nisus, la possibilité d'insérer du texte automatique se fait via un Glossaire, disponible pour les documents quel que soit le modèle par ailleurs utilisé. Ce Glossaire est un document contenant le texte cunéiforme et les entrées associées. Pour plus d'informations, voir le User Guide de Nisus Writer Pro.

Format: Cuneiform.ngloss

Installation:

User/Bibliothèque/Application Support/Nisus/Glossaries

il faut éventuellement importer le glossaire via le menu Préférences → Quickfix
→ Importer Glossaire (choisir Cuneiform.ngloss).

Utilisation du glossaire

taper l'entrée associée au caractère

taper "pomme d" pour développer l'entrée de glossaire en caractère cunéiforme.

NB. l'utilisation de texte automatique est grandement facilitée si on sélectionne dans les Préférences de Nisus **"étendre les entrées du glossaire lorsque vous tapez"**: taper "espace" donnera automatiquement le cunéiforme correspondant.

Pour ajouter, supprimer ou modifier des entrées de glossaire, ouvrir Cuneiform.ngloss et procéder aux changements. Chaque entrée de glossaire doit être séparée par un Séparateur d'entrée de glossaire (Menu Insertion). Le document doit être enregistré.

Sylvie Vanséveren

Chercheuse qualifiée F.R.S.-F.N.R.S.

Université Libre de Bruxelles

svsevere@ulb.ac.be